Theoretical Computer Science 47 (1986) 159-168
North-Holland

SUBWORD TOPOLOGY

V. Rajkumar DARE and Rani SIROMONEY
Department of Mathematics, Madras Christian College, Tambaram, Madras 600 059, India

Communicated by M. Nivat
Received January 1985
Revised May 1986

Abstract. We introduce a topology on a language $L \subset X^{\infty}$, called subword topology, which reflects certain interesting properties of subwords in a language. Well-known topological concepts such as compactness, closure of a language and closed sets reflect certain characteristic properties of subwords. The concept of adherence of a language (Nivat, 1979) is generalized to that of subword adherence and comparison is made with other limiting processes. The study of closed sets throws further light on Ehrenfeucht's Conjecture (Choffrut and Culik II, 1984) on unavoidable sets and generalizes it.

Introduction

A study on subwords of a language plays an important role in formal languages, especially in the developmental systems. Subword complexities of different kinds of D0L languages have extensively been studied in [3] (see also references cited therein). In this paper we study properties of subwords of a language using topological concepts.

A topology called SW topology is defined on a language $L \subset V^{\infty}$, where the sets $S_x = \{u \in L : x \text{ is a subword of } u\}$ are taken as a subbase. In this topology, every homomorphism is shown to be continuous. In general, languages need not be compact and Hausdorff under this topology. It is interesting to note that the well-known concepts of compact sets, closed sets and closure reflect interesting properties of subwords related to languages.

Compactness in SW topology reflects the following characterization regarding subwords in a language. A language $L \subset V^*$ is compact iff there exists a finite set $F \subset L$ such that every word in L has at least one member of F as its subword. Certain closure properties of compact languages are established.

A closed language in this topology reflects the fact that the set contains all the subwords of its elements and this plays an important role, too. Ehrenfeucht's Conjecture regarding unavoidable sets is that every unavoidable set L of V^* is extendible in the sense that there exists an $a \in V$ such that $L - \{x\} \cup \{xa\}$ itself is

0304-3975/86/\$3.50

unavoidable [1], where $x \in V^*$. We extend the concept of unavoidable sets to languages which are subsets of V^* and generalize Ehrenfeucht's Conjecture to closed languages. Certain results concerning this generalized conjecture are established which may throw light on the original conjecture.

The familiar topological concept of closure leads us to define the notion of S-adherence (subword adherence) similar to the concept of adherence introduced by Nivat [6]. Thus, we have one more device for generating infinite words from finite words. We have studied the algebraic properties of S-adherence and compared it with other well-known limiting processes. We have also established certain decidability results.

1. Subword topology

In this section we introduce a few topologies based on the structure of subwords of words in a language. We examine a few simple consequences of this definition which reflect significant properties relating to the structure of the language, corresponding to the well-known topological concept of compactness.

We recall that an infinite word u is a mapping $u:\{1,2,\ldots\}\to V$, where V is a finite alphabet and V^ω denotes the set of all such mappings. We write $u = a_1a_2\ldots$, where $u(n) = a_n$ for $n = 1, 2, \ldots$. Also, $V^\infty = V^* \cup V^\omega$, where V^* stands for the set of finite words over V.

If $x \in V^*$ and $y \in V^\infty$, then x is a subword of y iff $y = uxv$, where $u \in V^*$ and $v \in V^\infty$. Sub(y) denotes the set of all subwords of y and, for $L \subset V^\infty$, $\text{Sub}(L) = \bigcup_{y\in L}\text{Sub}(y)$. If $u = \lambda$, then x is a left factor of y and if $v = \lambda$, then x is a suffix of y. The set of all suffixes of a word y is denoted by Suff(y) and $\text{Suff}(L) = \bigcup_{y\in L}\text{Suff}(y)$. We define three topologies using the concept of subwords; the first is based on prefixes of words, the second on subwords of words, and the third on shuffle products on words. Let $L \subset V^\infty$.

(a) Let $u \in V^*$ and $P_u = \{x \in L$ such that $x = ux'$ where $x' \in V^*\}$. The topology generated by $\{P_u: u \in V^*\}$ in L is called PF-topology (prefix topology). We note that this coincides with the metric topology defined in [6].

(b) Let $u \in V^*$ and $S_u = \{x \in L: x = x'ux$ where $x', x \in V^*\}$. The topology generated by $\{S_u: u \in V^*\}$ in L is called SW-topology (subword topology).

(c) Let $u \in V^*$ and $\text{SF}_u = \{x \in L$ such that $x = x_1u_1x_2u_2\ldots x_nu_nx_{n+1}$, where u_i, $x_i \in V^*$ and $u = u_1u_2\ldots u_n\}$. The topology generated by $\{\text{SF}_u: u \in V^*\}$ in L is called SF-topology (shuffle topology).

Since, for each $u \in V^*$, $P_u \subset S_u \subset \text{SF}_u$, the following hierarchy exists for the above topologies: SF $\subset$ SW $\subset$ PF.

Since SW-topology reflects the properties of subwords, we consider it as the subject of our study.

A few simple consequences of the concept of SW-topology are listed in the following theorem.

Theorem 1.1. (a) *In general, there are languages which are not compact and not Hausdorff.*

(b) *Homomorphism is a continuous function in the subword topology.*

Proof. (a) The language $L=\{ab^na: n \geq 1\}$ is not compact and $L=\{a^nb^n: n \geq 1\}$ is not Hausdorff.

(b) Let $h: V_1^* \to V_2^*$ be a homomorphism. We now prove that $h: V_1^* \to V_2^*$ is continuous by showing that $h^{-1}(S_u)$ is open for $u \in V_2^*$. If $y \in h^{-1}(S_u)$, then $h(y) = xuz$, where $x \in V_2^*$ and $z \in V_2^*$. Since h is a homomorphism, there exists a subword v of y such that $y = x'vz'$ and $h(v) = v'uv''$. Then, $y \in S_v \subset h^{-1}(S_u)$ and hence $h^{-1}(S_u)$ is open. □

Note. The homomorphism $h: V_1^* \to V_2^*$ can be extended to a function $H: V_1^\infty \to V_2^\infty$ in the following sense: $H(u) = \lim\{h[u[n]]\}$, where $u \in V_1^\infty$. We note that this function is not a homomorphism. This mapping H is also a continuous function.

Since there are languages which are not compact, it is of interest to introduce the compact languages and study their closure properties. It is well known that a topological space is compact iff every open cover has a finite subcover [5]. We define a language $L \subset V^\infty$ to be compact iff it is compact in the subword topology.

We now give a characterization for a language in V^* to be compact.

Theorem 1.2. *A language $L \subset V^*$ is compact iff there exists a finite set $F \subset L$ such that every word in L has at least one member of F as its subword.*

Proof. If L is compact, then every open cover has a finite subcover. For the open cover $\{S_u: u \in L\}$ let $\{S_{u_i}: i = 1, 2, \ldots, n\}$ be a finite subcover. Denote $\{u_1, u_2, \ldots, u_n\}$ by F. If $x \in L$, it is in S_{u_i} for some i and hence u_i is a subword of x.

On the other hand, let $\{u_1, u_2, \ldots, u_n\}$ be a finite subset of L such that every word in L has at least one member of F as its subword.

If $\{S_\alpha\}$ is an open cover, then $u_i \in S_{\alpha_i}$ for some α_i. Then, clearly, $\{S_{\alpha_i}\}_{i=1}^n$ is the required finite subcover and hence L is compact. □

Hereafter we shall say that F is a finite subword set for the compact language L. The following corollaries are immediate and useful.

Corollary 1.3. *V^* and V^+ are compact.*

Corollary 1.4. *The Dyck set is compact and hence every homomorphic image of Dyck set is compact.*

We now examine the closure properties of compact languages.

Theorem 1.5. *The family of all compact languages in V^* is closed under union, homomorphism, catenation closure and inverse of λ-free homomorphism. But it is not closed under concatenation and intersection with regular sets.*

Proof. Since the finite union of compact sets and continuous image of compact sets are compact, we see that compact languages are closed under union and homomorphism.

Consider a compact language L with $\{u^{(1)}, u^{(2)}, \ldots, u^{(n)}\}$ as its finite subword set. Let $\{S_\alpha\}$ be an open cover for L^+. If $u^{(i)} \in S_{\alpha^{(i)}}$ for some $\alpha^{(i)}$, then $\{\alpha^{(1)}, \alpha^{(2)}, \ldots, \alpha^{(n)}\}$ is the finite subword set for L^+. Hence L^+ is compact. L^* is compact since it contains λ and the finite subword set for L^* is $\{\lambda\}$.

Now we prove that if L is a compact language and if h is a λ-free homomorphism, then $h^{-1}(L)$ is compact. Let $\{u^{(1)}, u^{(2)}, \ldots, u^{(n)}\}$ be a finite subword set for the compact set L. Let

$$A = \{v : u^{(i)} \text{ is a subword of } h(v) \text{ and there exists no subword } w \text{ of } v \text{ such that } u^{(i)} \text{ is a subword of } h(w) \text{ for some } i\}.$$

Since h is a λ-free homomorphism, this set A is finite. If $u \in h^{-1}(L)$, then $h(u) \in L$. So, $u^{(i)}$ is a subword of $h(u)$ for some i. Then, by the construction of the set A, there exists a $v \in A$ such that v is a subword of u. Hence, h^{-1} (L) is compact and A is the finite subword set corresponding to it. □

We note that V^* is compact and $L = \{ab^na : n \geq 1\}$ is regular but $V^* \cap L = L$ is not compact with $V = \{a, b\}$. Thus, compact languages are not closed under intersection with regular sets.

It is clear that the languages $L_1 = \{a\}$ and $L_2 = \{b^na : n \geq 1\}$ are compact languages but $L_1L_2 = \{ab^na : n \geq 1\}$ is not compact. Hence, compact languages are not closed under concatenation.

2. Closed sets and Ehrenfeucht's Conjecture

In this section we show that closed sets in this SW-topology play an important role in generalizing Ehrenfeucht's Conjecture on unavoidable sets and extendibility. Unavoidable sets in V^* have been studied in [1] and here we extend the concept to any language $L \subset V^*$. These studies may throw further light on Ehrenfeucht's Conjecture regarding unavoidable sets and we generalize Ehrenfeucht's Conjecture.

A word x is called a limit point of a language L iff every neighbourhood of that point has a word in L other than x. That is, L has at least one word having x as its subword. A language L is said to be a closed language iff it contains all its limit points. It follows that the set of all subwords of L forms the set of all limit points of L. Thus, L is a closed language iff $L = \mathrm{Sub}(L)$.

Example 2.1. If $V = \{a, b\}$, then $L = a^*b^*$ is a closed language because it contains all its subwords. But $L = \{a^nb^n : n \geq 1\}$ is not closed since subwords of the form a^ib^j are not in L for $i \neq j$.

Definition 2.2. A set X of a language L is said to be *unavoidable* in L iff all but a finite number of words in L has a subword in X.

An unavoidable set in L is said to be *minimal* iff no subset of X is unavoidable in L.

Example 2.3. For the language $L=\{a^nb^na^n, ab^na: n\geqslant 1\}$, $X=\{ab^na: n\geqslant 1\}$ is a minimal unavoidable set and $|X|=\infty$. We note that this result is in contrast with the one given in [1] that every minimal unavoidable set in V^* is finite.

Definition 2.4. An element x of X is *extendible* if $X-\{x\}\cup\{y\}$ is unavoidable in L, where $y\in S_x\cap L$. An unavoidable set is extendible if it has at least one extendible element.

We note that every unavoidable set of a language need not be extendible. For example, if $L=\{ab, ab^na: n\geqslant 1\}$, then $X=\{ab\}$ is an unavoidable set which is not extendible.

Generalized Ehrenfeucht Conjecture. *Every unavoidable set in a closed language is extendible.*

This generalizes Ehrenfeucht's Conjecture which states that "every unavoidable set of V^* is extendible".

We first give a characterization of extendibility of an unavoidable set in a language L.

Theorem 2.5. *Let $X\subset L$ be an unavoidable set of L. Then $x\in X$ is extendible iff there exists a $y\in S_x\cap L$ such that $L-\bigcup_{z\in X-\{x\}} S_z - S_y$ is finite.*

The proof is clear from the definition.

The following results are immediate from the definition:

(1) A language L is compact iff every minimal unavoidable set of L is finite.

(2) If $L^{\geqslant n}$ is the set of all words of L of length greater than or equal to n, then X is unavoidable iff there exists an $n>0$ such that $L^{\geqslant n}\subset V^*XV^*$.

(3) If $Y\subset V^*XV^*\cap L$ is unavoidable in L, then X is also unavoidable.

(4) The following two conditions are equivalent:

(i) For every unavoidable set X of L there is an $x\in X$ and a $y\in S_x\cap L$ such that $X-\{x\}\cup\{y\}$ is unavoidable.

(ii) For every unavoidable set $X\subset L$ there exists an increasing sequence of words $\{y^{(n)}\}$ in S_x such that $X_n=X-\{x\}\cup\{y^{(n)}\}$ is unavoidable in L.

The following theorem shows that extendibility is unique.

Theorem 2.6. *Let X be a minimal unavoidable set of a language L and $x\in X$. If $X_1=X-\{x\}\cup\{y\}$ is unavoidable, there exists no $y'\in S_x$ which is incomparable with y such that $X_2=X-\{x\}\cup\{y'\}$ is unavoidable in L.*

Proof. Since X_1 is unavoidable, $L - \bigcup_{z \in X - \{x\}} S_z - S_y$ is finite. Similarly, if X_2 is unavoidable, $L - \bigcup_{z \in X - \{x\}} S_z - S_{y'}$ is finite. Since y and y' are incomparable, we have a contradiction. $\square$

We now show that unavoidability and extendibility are preserved under continuous mappings.

Theorem 2.7. *Let $X \subset L$ be a finite unavoidable set of a language L. If f is a continuous function, then $f(x)$ is unavoidable in $f(L)$.*

Proof. Since X is unavoidable in L, $L - \bigcup_{x \in X} S_x$ is finite. This implies that $f(L - \bigcup_{x \in X} S_x)$ is finite.

But $f(L) - \bigcup_{x \in X} S_{f(x)} \subset f(L - \bigcup_{x \in X} S_x)$ since f is continuous. Hence, $f(L) - \bigcup_{x \in X} S_{f(x)}$ is finite, which implies that $f(X)$ is unavoidable in $f(L)$. $\square$

Theorem 2.8. *If $x \in X$ is extendible, then $f(x)$ in $f(X)$ is extendible.*

The proof follows from the fact that if f is continuous, then x is a subword of y implies that $f(x)$ is a subword of $f(y)$.

The generalization of Ehrenfeucht's Conjecture and the results established here may throw further light in proving Ehrenfeucht's Conjecture.

3. Closure and S-adherence

In this section we study certain aspects of closure of a language and subword adherence.

In a topological space X, the closure of a set L is defined as $\bar{L} = \{x \in X: \text{every neighbourhood of } x \text{ intersects } L\}$.

Extending the concept of adherence introduced by Nivat [6] we now introduce subword adherence of a language L, called S-adherence of L, and examine its properties, relating it to earlier known concepts of limiting processes. Thus, we have one more device for obtaining infinite words from finite words. For a language L, this generalizes the definition of adh(L) since, for each $L \subset V^\infty$, adh(L) is contained in S-adh(L).

Definition 3.1. For a language $L \subset V^\infty$,

$$\text{S-adh}(L) = \{x \in V^\infty : \text{Sub}(x) \subset \text{Sub}(L)\}.$$

Theorem 3.2. *If $L \subset V^\infty$, then $\bar{L} = \text{Sub}(L) \cup \text{S-adh}(L)$.*

Proof. If we consider a word x in $\bar{L} \cap V^*$, then $S_x \cap L \neq \emptyset$. This implies that there exists a word in L having x as its subword. Thus, $x \in \text{Sub}(L)$.

On the other hand, if we consider x in $\bar{L} \cap V^\omega$ with $y \in \text{Sub}(x)$, then $x \in S_y$.

Since S_y is a neighbourhood for x, there exists a word z in L such that $z \in S_y$. Then, $y \in \text{Sub}(z) \subset \text{Sub}(L)$. Thus, we have $\text{Sub}(x) \subset \text{Sub}(L)$, hence $x \in \text{S-adh}(L)$. Therefore, $\bar{L} \subset \text{Sub}(L) \cup \text{S-adh}(L)$.

If we consider a word x in S-adh(L), then $\text{Sub}(x) \subset \text{Sub}(L)$. So, every neighbourhood of x intersects L and hence S-adh$(L) \subset \bar{L}$. Similarly, if $x \in \text{Sub}(L)$, then $x \in \bar{L}$. Therefore, $\text{Sub}(L) \cup \text{S-adh}(L) \subset \bar{L}$. □

We recall the definition of adherence [6] and limit [2] of a language L.

For any language L,

$$\text{Adh}(L) = \{x \in V^\omega : \text{FG}(x) \subset \text{FG}(L)\},$$

$$\text{Lim}(L) = \{x \in V^\omega : \text{for each } n \in N, \text{ there exists an integer } k \geqslant n \text{ such that } x[k] \in L\}.$$

We now examine some properties of S-adherence of a language L.

Theorem 3.3. *For $L \subset V^\infty$, we have*

(a) $\text{Lim}(L) \subset \text{adh}(L) \subset \text{S-adh}(L)$,

(b) $\text{S-adh}(L) = \text{S-adh}(\text{Sub}(L)) = \text{adh}(\text{Sub}(L)) = \lim(\text{Sub}(L))$,

(c) $\text{S-adh}(L_1 \cup L_2) = \text{S-adh}(L_1) \cup \text{S-adh}(L_2)$,

(d) $\text{S-adh}(L^*) = \text{Suff}(L)L^\omega \cup L^*\text{Adh}(L)$.

The proof directly follows from the definition of S-adherence and by simple computation.

We now compare the family of S-adherences with the family of adherences and the family of limit languages.

Theorem 3.4. *If* S-ADH *is the family of all* S-*adherences,* ADH *the family of all adherences and* LIM *the family of all limit languages, then* $\text{S-ADH} \subsetneq \text{ADH} \subsetneq \text{LIM}$.

Proof. If L is in S-ADH, then there is an $L_1 \subset V^\infty$ such that $L = \text{S-adh}(L_1)$. But,

$$\text{S-adh}(L_1) = \text{S-adh}(\text{Sub}(L_1)) = \text{adh}(\text{Sub}(L_1))$$

and hence $L \in \text{ADH}$.

Similarly, we prove that $\text{ADH} \subset \text{LIM}$. We note that $ab^\omega \in \text{ADH}$, but not in S-ADH. This follows from the fact that $\text{S-adh}(L) = \text{S-adh}(\text{Sub}(L))$. Also, $a^*b^\omega \in \text{LIM}$, but not in ADH [2]. Thus $\text{ADH} \subsetneq \text{LIM}$. □

We now compare the family of S-adherences of D0L systems and languages in Chomskian hierarchy with families obtained from other limiting processes.

Definition 3.5. Let $\mathscr{L}$ be a family of languages. Then,

$$\text{S-ADH}(\mathscr{L}) = \{L : L = \text{S-adh}(L_1), \text{ where } L_1 \in \mathscr{L}\}.$$

Theorem 3.6. *For the family* D0L, S-ADH(D0L) *is incomparable with* LIM(D0L) *and* ADH(D0L).

Proof. The language $\{ab^\omega\}$ is in LIM(D0L) but not in S-ADH(D0L). If $G = \{(a, b), (a \to ab, b \to ba), a\}$, then $L_1 = \text{S-adh}(L(G))$ is infinite. But, limits and adherences of D0L are finite. Therefore, $L_1 \notin \text{LIM(D0L)}$ and ADH(D0L). Hence, we find that S-ADH(D0L) is incomparable with LIM(D0L) and ADH(D0L). □

A family of languages is called a *trio* if it is closed under the three operations of λ-free homomorphism, inverse homomorphism and intersection with regular sets. A trio is said to be a *full trio* if it is closed under arbitrary homomorphism. It has been shown that if $\mathscr{L}$ is a full trio and $L \in \mathscr{L}$, then $\text{Sub}(L) \in \mathscr{L}$ [4]. We denote the family of recursively enumerable languages, context-free languages, linear languages and regular languages by RE, CF, LIN, and REG respectively. It is well known that these language families are full trios [4]. Whereas for the D0L family, S-ADH(D0L), ADH(D0L), and LIM(D0L) are incomparable, for the language families RE, CF, LIN, and REG there is proper containment as can be seen from the following theorem.

Theorem 3.7

$$\text{S-ADH}(\mathscr{L}) \subsetneq \text{ADH}(\mathscr{L}) \subsetneq \text{LIM}(\mathscr{L}) \quad \textit{if } \mathscr{L} \in \{\text{RE, CF, LIN, REG}\}.$$

Proof. Let $L \in \text{S-ADH}(\mathscr{L})$. Then $L = \text{S-Adh}(L_1)$, where $L_1 \in \mathscr{L}$. Since $\mathscr{L}$ is a full trio, $\text{Sub}(L_1) \in \mathscr{L}$. But, $\text{S-adh}(L_1) = \text{S-adh}(\text{Sub}(L_1)) = \text{lim}(\text{Sub}(L_1))$. Thus, we have $\text{S-ADH}(\mathscr{L}) \subsetneq \text{ADH}(\mathscr{L}) \subsetneq \text{LIM}(\mathscr{L})$. □

We now establish certain decidability results.

Theorem 3.8. *For an arbitrary language L, a word u is a subword of an infinite word in* S-adh(L) *iff it is a prefix of infinitely many subwords of L.*

The proof follows from the fact that $\text{S-adh}(L) = \text{lim}(\text{Sub}(L))$.

The following result shows that a D0L language can be decomposed into a finite number of D0L systems such that the S-adherence of the former is the union of a finite set of S-adherences of the latter. The decomposed D0L languages are simple in that the axiom is a single letter of the alphabet and this result is useful in proving certain decidability results.

Theorem 3.9. *Let* $G = (V, h, w)$ *be an infinite* D0L *system and* $A = \{a \in V : |L(G_a)| = \infty$ *where* $G_a = (V, h, a)\}$.

Then, $\text{S-adh}(L(G)) = \bigcup_{a \in A} \text{S-adh}(L(G_a))$.

Proof. Let $u \in \text{S-adh}(L(G))$, $u = \lim\{\alpha^{(i)}\}$, where $\alpha^{(i)} \in \text{Sub}(L(G))$. Choose $a^{(i)}$ from A such that $\alpha^{(i)} \in \text{Sub}(L(G_{a^{(i)}}))$, $i = 1, 2, \ldots$. Since A is finite, there is at least one element in $\{a^{(i)}: i = 1, 2, \ldots\}$ which must repeat an infinite number of times. Let this element be $a^{(j)}$; then, $u \in \text{S-adh}(L(G_{a^{(j)}}))$ implies $u \in \bigcup \text{S-adh}(L(G_{a^{(j)}}))$. Thus, we have $\text{S-adh}(L(G)) = \bigcup_{a \in A} \text{S-adh}(L(G_a))$. □

Remark. It is interesting to note that the corresponding result is not true for adherences as can be seen from the following example: for, if $G = \{(a, b), (a \to aa, b \to bb), ab\}$, then $A = \{a, b\}$. But $\text{adh}(L(G)) = a^{\omega}$, $\text{adh}(L(G_a)) = a^{\omega}$, and $\text{adh}(L(G_b)) = b^{\omega}$.

Theorem 3.10. *If L and L' are two languages, then* $\text{S-adh}(L) = \text{S-adh}(L')$ *implies that for each* $u \in \text{S-adh}(L)$ *there exists the set*

$$\{u^{(1)}, u^{(2)}, \ldots : u^{(i)} \textit{ is a prefix of } u^{(i+1)}\} \subset \text{Sub}(L) \cap \text{Sub}(L')$$

such that $u = \lim\{u^{(1)}, u^{(2)}, \ldots\}$.

Proof. Since $\text{S-adh}(L) = \lim(\text{Sub}(L))$, if $u \in \text{S-adh}(L) = \text{S-adh}(L')$, then $u = \lim\{u^{(i)}\} = \lim\{v^{(j)}\}$, where $u^{(i)} \in \text{Sub}(L)$ and $v^{(j)} \in \text{Sub}(L')$. For each i and j we have either $u^{(i)} \leqslant v^{(j)}$ or $v^{(j)} \leqslant u^{(i)}$. Thus, we have a common sequence for u. □

Theorem 3.11. *For any arbitrary languages L_1 and L_2,* $\text{S-adh}(L_1) \subset \text{S-adh}(L_2)$ *iff every infinite increasing sequence of elements in* $\text{Sub}(L_1)$ *is also in* $\text{Sub}(L_2)$.

The proof of the theorem follows from the definition of S-adherence.

Now we prove the decidability of the S-adherence equivalence problem for D0L languages.

Theorem 3.12. *Let $G_1 = (V, h, a_1)$ and $G_2 = (V, g, a_2)$ be two* D0L *systems, where $a_i \in V$. Then,* $\text{S-adh}(L(G_1)) \subset \text{S-adh}(L(G_2))$ *iff the following condition is satisfied: If $u_i \in \text{Sub}(h^{n_i}(a_i))$, $i = 1, 2$, such that $u_1 < u_2$, then for the pair of integers (n_1, n_2) there exists another pair of integers (m_1, m_2) such that $u_i \in \text{Sub}(h^{n_i}(a_i)) \cap \text{Sub}(g^{m_i}(a_2))$.*

Proof. If $\text{S-adh}(L(G_1)) \subset \text{S-adh}(L(G_2))$, then, by Theorem 3.10, for each $u \in \text{S-adh}(L(G_1))$, there exists a common increasing sequence of subwords of $L(G_1)$ and $L(G_2)$. Thus, when $u_1 \leqslant u_2$ for $u_i \in \text{Sub}(h^{n_i}(a_i))$, there exist integers m_1 and m_2 such that $u_i \in \text{Sub}(g^{m_i}(a_2))$.

Conversely, we assume that the given condition is satisfied. Since $u_1 \in \text{Sub}(h^{n_1}(a_1))$ and $u_2 \in \text{Sub}(h^{n_2}(a_1))$, we have $u_2 \in \text{Sub}(h^{n_2-n_1}(u_1'))$. Similarly, $u_2 \in \text{Sub}(g^{m_2-m_1}(u_1''))$, where $u_1' = h^{n_1}(a_1)$ and $u_1'' = g^{m_1}(a_2)$.

By the property of parallel rewriting there exists an $u_2 < u_3$ such that $u_3 \in \text{Sub}(h^{n_2-n_1}(u_2')) \cap \text{Sub}(g^{m_2-m_1}(u_2''))$, where $u_2' = h^{n_2}(a_1)$ and $u_2'' = g^{m_2}(a_2)$.

Proceeding in this way we get an increasing sequence $u_1 < u_2 < \cdots$ which is common to both $\text{Sub}(L(G_1))$ and $\text{Sub}(L(G_2))$. Hence, $\text{S-adh}(L(G_1)) \subset \text{S-adh}(L(G_2))$. □

The concept of S-adherence may be used to generate bi-infinite words.

Let ${}^{\omega}V^{\omega}$ be the set of all bi-infinite words on V. We define the relation $<_s$ in V^* such that, for $x, y \in V^*$, $x <_s y$ iff $y = x'xx''$, where $x', x'' \in V^+$. Let $\{x^{(i)}\}$ be a sequence in V^* such that $x^{(i)} <_s x^{(i+1)}$. Then the limit of this sequence with the ordering $<_s$ is a bi-infinite word and we denote it by bi-lim$\{x^{(i)}\}$. For $L \subset V^*$, we define bi-lim(L) as the collection of all bi-limits of the sequences of the above form.

Then, we have

$$\text{S-Bi-adherence}(L) = \{u \in {}^{\omega}V^{\omega} : \text{Sub}(u) \subset \text{Sub}(L)\} = \text{Bi-lim}\{\text{Sub}(L)\}.$$

Thus, we have seen that well-known topological concepts such as compactness, closedness and closure reflect interesting properties of subwords related to languages.

References

[1] C. Choffrut and K. Culik, II, On extendibility of unavoidable sets, Lecture Notes in Computer Science **166** (Springer, Berlin, 1984) 326–338.

[2] K. Culik, II and A. Salomaa, On infinite words obtained by iterating morphisms, *Theoret. Comput. Sci.* **19** (1981) 29–38.

[3] A. Ehrenfeucht and G. Rozenberg, On the subword complexity of locally catenative DOL languages, *Inform. Process. Lett.* **16** (1983) 7–9.

[4] S. Ginsburg, *Algebraic and Automata-Theoretic Properties of Formal Languages* (North-Holland, Amsterdam, 1975).

[5] J.L. Kelly, *General Topology* (Van Nostrand, Princeton, NJ, 1955).

[6] M. Nivat, Infinite words, infinite trees and infinite computations, *Mathematical Centre Tracts* **109** (1979) 1–52.